A Appendix: Datasheet

A.1 Motivation

- **For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled?** This dataset is intended to be used for training neural networks for reconstructing depth maps and/or surface normal maps from a single RGB image. It contains objects without any textures on them, and is specifically intended for training networks for texture-less surface reconstruction. It was created because of a lack of any existing large scale datasets for this task.

- **Who created the dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?** It was created by a research assistant in the Augmented Vision department at German Research Center for Artificial Intelligence (DFKI).

- **Who funded the creation of the dataset?** This project was not funded.

A.2 Composition

- **What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)?** Each instance in the dataset consists of an RGB image (.png) of one texture-less object, and a compressed NumPy array file (.npz) containing the corresponding depth map and a surface normals map for the image. 35 different objects are featured in the dataset, including some clothing items, animals, furniture, vehicles, statues and some other miscellaneous items.

- **How many instances are there in total (of each type, if appropriate)?** There are 8640 instances per object and 302,400 instances in total.

- **Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set?** This dataset represents a subset of a non-exhaustive set of all texture-less objects that exist in the real-world. To ensure sufficient coverage, the dataset contains a diverse collection of objects which represent objects of different scales and geometrical complexities. Each object is modeled under several lighting conditions and rendered from different cameras looking at the object from various perspectives. This ensures that as many lighting and perspective variations are represented in the dataset as possible.

- **What data does each instance consist of?** Each instance consists of an image of a 3D object rendered under realistic viewing conditions, and the 3D information of the parts of the object visible in the image that was generated by a 3D modeling software. Both the images and 3D information are raw data.

- **Is there a label or target associated with each instance?** Yes, for each RGB image, there are two labels provided: a depth map and a surface normals map. Both of these labels are stored in a single compressed NumPy file.

- **Is any information missing from individual instances?** No, there is no information missing for any of the objects or instances.

- **Are relationships between individual instances made explicit (e.g., users’ movie ratings, social network links)?** The directory structure and file names make the relationships between images and labels, as well as different objects, the lighting conditions and camera angles explicit.

- **Are there recommended data splits (e.g., training, development/validation, testing)?** For each object, we recommend using the sequences where all lights were on (named “La_*)” for testing, and using all the other sequences (where individual lights were turned on one by one) for training. This gives a train/test split ratio of 75 to 25. This split is recommended because it would allow the network to learn features that are independent of shadows and lighting.

- **Are there any errors, sources of noise, or redundancies in the dataset?** No.

- **Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)?** The dataset does not rely on any external resources.
• Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor–patient confidentiality, data that includes the content of individuals’ non-public communications)? No.

• Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? No.

A.3 Collection Process

• How was the data associated with each instance acquired? The images and depth maps were directly generated by the 3D modeling software, Blender. Surface normals were computed from the depth maps by differentiating them, and visualized as images to manually verify their correctness.

• What mechanisms or procedures were used to collect the data (e.g., hardware apparatuses or sensors, manual human curation, software programs, software APIs)? A popular open-source 3D modeling software Blender was used to generate the data. Details of the generation process are provided in the paper, and the source code used for data generation is provided on GitHub.

• If the dataset is a sample from a larger set, what was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)? There was no defined sampling strategy, as it is difficult to define the larger set itself. We included the 3D models that are popularly used in Computer Graphics and Vision applications, such as Stanford Bunny and Utah teapot among others. In addition to that, we included models which are similar to objects in existing texture-less datasets (particularly the clothing items). Finally, we also included objects based on their physical size in real world such that there are objects ranging from a rubber duck to a spaceship and everything in between.

• Who was involved in the data collection process and how were they compensated? The dataset was collected by a student working on this project as part of their Master’s thesis.

• Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (e.g., recent crawl of old news articles)? The dataset was created between September 2021 and March 2022.

• Were any ethical review processes conducted (e.g., by an institutional review board)? No.

A.4 Preprocessing/Cleaning/Labeling

• Was any preprocessing/cleaning/labeling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? No, there was no preprocessing.

A.5 Hosting and License

The primary dataset can be viewed and downloaded through its homepage at https://projects.dfkI.uni-kl.de/textureless_object_data/. This webpage is owned by theTechnical University of Kaiserslautern and German Research Center for Artificial Intelligence (DFKI). DFKI ensures that the website remains available in the long term.

The homepage provides curated access to all files in the dataset, a link to the supplementary dataset (which is hosted publicly by the author on a Kaggle repository), links to all associated source codes, research paper, and any additional material related to this dataset. Data loaders for reading the dataset using PyTorch framework in Python are also available in the source code repository linked on the homepage.

The dataset is provided under the CC Attribution 4.0 International (CC BY 4.0) license, and all source code is provided with the MIT License, unless stated otherwise in the code repository. The source code is hosted publicly on GitHub.